

QUADRATIC COVARIATION ESTIMATES IN NONSMOOTH STOCHASTIC CALCULUS

SERGIO ANGEL ALMADA MONTER AND YURI BAKHTIN

ABSTRACT. Given a Brownian Motion W , in this paper we study the asymptotic behavior, as $\varepsilon \rightarrow 0$, of the quadratic covariation between $f(\varepsilon W)$ and W in the case in which f is not smooth. We use a recent representation as a backward- forward Itô integral of $[f(\varepsilon W), W]$ to prove an ε -dependent approximation scheme which is of independent interest. We get the result by providing estimates to this approximation.

Non-smooth Itô's formula and Quadratic Variation and Large Deviation

1. INTRODUCTION

One of the central results of stochastic calculus is Itô's change of variables formula for twice differentiable transformations of semimartingales. It was realized recently that one also needs to study nonlinear maps that are not smooth enough to allow an application of the classical Itô formula. Various approaches to less regular changes of variables have been introduced, see [3], [4], [6], [8], [14], and references therein. These studies show that the key feature of the Itô formula, the quadratic covariation term, is well-defined under much weaker assumptions than those leading to the traditional formula. However, no nontrivial quantitative estimates of the arising quadratic covariation processes have appeared in the literature, to the best of our knowledge.

One area where such estimates are naturally needed is small random perturbations of dynamical systems. Often, in the course of a study of a stochastic system one has to make a simplifying change of coordinates, transforming the system locally to a simpler one. If the transformation map is C^2 , then one can apply the classical Itô calculus and easily control the Itô correction term. However, there are situations where a natural change of variables is less regular than C^2 , and in these cases there is no readily available tool that could be used to control the generalized Itô correction.

The goal of this paper is to close this gap and provide quantitative estimates on the generalized Itô correction term under non classical assumptions on the transformation.

Let us now be more precise. Let W be a standard 1-dimensional Wiener process on a complete probability space (Ω, Σ, P) and $\varepsilon > 0$ be a constant. If $g : \mathbb{R} \rightarrow \mathbb{R}$ is C^2 , then the classical Itô formula is (see [13, Section II.7])

$$g(\varepsilon W(t)) - g(0) = \varepsilon \int_0^t g'(\varepsilon W(s)) dW(s) + \frac{\varepsilon^2}{2} \int_0^t g''(\varepsilon W(s)) ds.$$

Introducing $f = g' \in C^1$, we can also rewrite the second term in the r.h.s. as quadratic covariation between $f(\varepsilon W)$ and εW : for $Q_\varepsilon(t) = [f(\varepsilon W), \varepsilon W](t)$, we

have

$$Q_\varepsilon(t) = \varepsilon^2 \int_0^t f'(\varepsilon W(s)) ds, \quad t \geq 0.$$

In particular, for any $T > 0$, $\varepsilon^{-1} \sup_{t \leq T} Q_\varepsilon(t) \rightarrow 0$ in probability as $\varepsilon \rightarrow 0$. In this paper we show that this convergence holds in the case in which f is not differentiable.

The motivation for this problem relies on small random perturbation of dynamical systems. Suppose that b is a vector field with a critical point at x^* and let S denote the flow generated by b :

$$\frac{d}{dt} S^t x = b(S^t x), \quad S^0 x = x.$$

It is well known (see Section 2.8 of [12]) that there is an homeomorphism g so that locally around $g(x^*)$ the flow $g(S^t x)$ behaves like the linearized version of S , e^{At} , where A is the Jacobian of b at x^* . This idea combined with the traditional Itô's formula imply (see, e.g. [1], [11]) that if g is at least C^2 and x_0 is close enough to x^* , the system

$$dX_\varepsilon(t) = b(X_\varepsilon(t))dt + \varepsilon dW(t), \quad X_\varepsilon(0) = x_0,$$

could be analyzed by working with

$$d\tilde{X}_\varepsilon(t) = \left(A\tilde{X}_\varepsilon(t) + \frac{\varepsilon^2}{2} g''(X_\varepsilon(t)) \right) dt + \varepsilon g'(g^{-1}(\tilde{X}_\varepsilon(t))) dW(t), \quad \tilde{X}_\varepsilon(0) = g(x_0).$$

In this case, the analysis of \tilde{X} can be carried out as the analysis of an Ornstein-Uhlenbeck kind of process with a perturbation of order ε^2 , simplifying the problem considerably. Unfortunately, g is not in general C^2 ; instead, there is a well understood range of cases for which $g \in C^1$ (e.g. Hartman Theorem on Section 2.8 of [12]). In these cases, an already known C^1 formulation of Itô's formula implies a similar analogy between the non-linear and linear systems with $Q_\varepsilon^b = [g'(X_\varepsilon), X_\varepsilon]$ instead of $(\varepsilon^2/2) \int_0^t g''(X_\varepsilon(r)) dr$. Hence, an estimate that shows that in these cases $\varepsilon^{-1} Q_\varepsilon^b \rightarrow 0$, as $\varepsilon \rightarrow 0$, allows for the same simplification as in the case in which $g \in C^2$. Up to the authors knowledge, there are no estimates like this in the literature. The objective of this paper is to prove the simplest of these estimates; that is, that, f continuous, $\varepsilon^{-1} [f(\varepsilon W), \varepsilon W] \rightarrow 0$, as $\varepsilon \rightarrow 0$, in probability.

The analysis of the quadratic covariation $[g'(X), X]$ in connection with extensions of Itô's formula for functions $g \notin C^2$ is fundamental for nonsmooth Itô calculus, see [5], [7], [8], [14], [15]. In [8], [14], [15] methods from backward stochastic calculus were used (see also the summary [16]), while in [5], [7] a local time approach was used. The basic result that has been explained in the cited literature from several points of view is that for $T > 0$,

$$(1) \quad Q_\varepsilon(t) = -\varepsilon \int_0^t f(\varepsilon W(s)) dW(s) - \varepsilon \int_{T-t}^T f(\varepsilon W(T-s)) dW(T-s),$$

where both integrals can be understood as Itô integrals w.r.t. appropriate filtrations. It is well known [13, page 389] that the integral with respect to $W(T - \cdot)$ in (1) is the time reversal of a semimartingale w.r.t. the natural filtration of $W(T - \cdot)$. Here the time reversal (with respect to $T > 0$) of a process X is understood as $X(T - t) - X(T)$.

In this paper we exploit this structure by constructing an approximation scheme for Q_ε and using martingale techniques to show its consistency. As far as we know,

this is the first attempt to use such a scheme in small noise analysis. See [2] for a related but different scheme for local time approximation in the case $\varepsilon = 1$.

The text is organized as follows. In Section 2 we state our main results that include the martingale representation for the quadratic covariation. A proof of the martingale representation is given in Section 3. In Section 4 we use the martingale representation to propose an approximation scheme that we then use to prove the key bound that the main results depend upon. The proofs of the main theorems are given in Section 5. In Section 6 proofs of auxiliary lemmas are given.

2. MAIN RESULTS

We are going to study $Q_\varepsilon(t) = [f(\varepsilon W), \varepsilon W](t)$ assuming that $f : \mathbb{R} \rightarrow \mathbb{R}$ is a bounded and uniformly Hölder or Lipschitz function, although these assumptions on f can be relaxed. It is convenient to formulate these assumptions in terms of modulus of continuity defined by:

$$\text{osc}_f(\delta) = \sup_{|t-s| < \delta} |f(s) - f(t)|, \quad \delta > 0.$$

Throughout the text, we work with an arbitrary fixed number $T > 0$. We will not be explicit when including the dependency on $T > 0$ in the notation. We are ready to state the main results of the text.

Theorem 1. *Suppose $\text{osc}_f(\delta) \leq C_f \delta^\alpha$ for some $\alpha \in (0, 1)$, $C_f > 0$, and all sufficiently small δ . Then, for every $\delta > 0$, $\gamma \in (0, \alpha)$, and $\mu \in (\gamma, \alpha)$, there are constants $\varepsilon_{\delta, \mu} > 0$ and $C_{\delta, \mu} > 0$ such that*

$$\mathbb{P} \left\{ \varepsilon^{-(1+\gamma)} \sup_{t \leq T} |Q_\varepsilon(t)| > \delta \right\} \leq C_{\delta, \mu} \varepsilon^{2(\alpha-\mu)/(1-\alpha)}, \quad \varepsilon \in (0, \varepsilon_{\delta, \mu}).$$

In particular, for any $\gamma \in (0, \alpha)$,

$$\varepsilon^{-(1+\gamma)} \sup_{t \leq T} |Q_\varepsilon(t)| \xrightarrow{\mathbb{P}} 0, \quad \varepsilon \rightarrow 0.$$

This result is stronger than our initial claim that $\varepsilon^{-1} Q_\varepsilon \rightarrow 0$. Moreover, if α is close to 1, the exponent $1 + \gamma$ can be chosen to be close to 2.

The method we employ to prove this theorem produces the following estimate in the Lipschitz case where $\alpha = 1$:

Theorem 2. *Suppose $\text{osc}_f(\delta) \leq C_f \delta$, for some constant $C_f > 0$ and sufficiently small $\delta > 0$. Then, for every $\delta > 0$, $\gamma \in (0, 1)$, and $\mu \in (\gamma, \alpha)$, there are constants $\varepsilon_{\delta, \mu} > 0$ and $C_{\delta, \mu} > 0$ such that*

$$\mathbb{P} \left\{ \varepsilon^{-(1+\gamma)} \sup_{t \leq T} |Q_\varepsilon(t)| > \delta \right\} \leq C_{\delta, \mu} e^{-\varepsilon^{-(1-\mu)}}, \quad \varepsilon \in (0, \varepsilon_{\delta, \mu}).$$

This theorem establishes that the rate of decay in probability is exponential in the Lipschitz case, which is coherent with the differentiable case in which almost sure convergence holds. For the Hölder case, the method only shows a polynomial upper bound which in principle does not imply that the convergence rate can not be exponential.

The proof of Theorems 1 and 2 will be given in Section 5. An important part of the analysis is Theorem 7 given in Section 4 and in principle one can apply that result and its possible extensions to less regular functions f . The proof of

Theorem 7 is in turn based on a forward-backward martingale representation of the quadratic covariation that we proceed to explain.

Definition 3. *The time reversal of a process $X = (X(t))_{t \geq 0}$ with respect to $T > 0$ is defined by*

$$\bar{X}(t) = X(T - t) - X(T), \quad t \in [0, T].$$

Likewise, the backward of X with respect to $T > 0$ is defined by

$$\hat{X}(t) = X(T - t), \quad t \in [0, T].$$

The starting point is the representation for $L_\varepsilon = \varepsilon^{-1}Q_\varepsilon$ implied by (1). For any $T > 0$,

$$(2) \quad L_\varepsilon(t) = - \int_0^t f(\varepsilon W(s)) dW(s) - \int_{T-t}^T f(\varepsilon \hat{W}(s)) d\hat{W}(s), \quad t \in [0, T].$$

We will find a convenient way to rewrite this expression using an enlargement of filtration approach. Denoting the natural filtration of a process $X = (X_t)_{t \geq 0}$ by $\mathcal{F}^X = (\mathcal{F}_t^X)_{t \geq 0}$, we note that the integral with respect to W in (2) is an \mathcal{F}^W martingale, while the integral with respect to \hat{W} is the time reversal of the $\mathcal{F}^{\hat{W}}$ semimartingale

$$N_\varepsilon(t) = \int_0^t f(\varepsilon \hat{W}(s)) d\hat{W}(s).$$

Therefore, one of the terms in (2) is a martingale, while the other one has a non-trivial drift component. The following result reveals the structure of this time reversal.

Theorem 4. *Let $\mathcal{G} = (\mathcal{G}_t)_{t \in [0, T]}$ be the minimum filtration such that $W(T)$ is \mathcal{G}_0 measurable and $\mathcal{F}_t^{\hat{W}} \subset \mathcal{G}_t$. Then, \hat{W} is a \mathcal{G} semimartingale with Doob-Meyer decomposition given by*

$$(3) \quad \hat{W}(t) = W(T) - \int_0^t \frac{\hat{W}(s)}{T-s} ds + \beta(t),$$

for some Brownian Motion β with respect to \mathcal{G} .

Moreover, if $\mathcal{H} = (\mathcal{H}_t)_{t \in [0, T]}$ is the minimum complete filtration such that $W(T)$ is \mathcal{H}_0 measurable and $\mathcal{F}_t^{\hat{W}} \subset \mathcal{H}_t$, then β is an \mathcal{H} Brownian Motion, \hat{W} is an \mathcal{H} semimartingale with the Doob-Meyer decomposition (3) and \hat{W} can be written as

$$(4) \quad \hat{W}(t) = W(T)(1 - t/T) + (T - t) \int_0^t \frac{d\beta(s)}{T-s}, \quad t \in [0, T].$$

Remark 5. In particular, since \hat{W} is \mathcal{H} adapted, for every function $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ such that

$$\mathbb{E} \int_0^T F(\hat{W}(s))^2 ds < \infty,$$

the process $t \mapsto \int_{T-t}^T F(\hat{W}(s)) d\beta(s)$ is the time reversal of a martingale.

Using Theorem 4, we can obtain a representation for $L_\varepsilon = \varepsilon^{-1}Q_\varepsilon$. This is given in the following:

Corollary 6. *Let \mathcal{H} be as in Theorem (4). Then, the process $L_\varepsilon = \varepsilon^{-1}Q_\varepsilon$ can be written as*

$$(5) \quad L_\varepsilon(t) = - \int_0^t f(\varepsilon W(s)) dW(s) - \int_{T-t}^T f(\varepsilon \hat{W}(s)) d\beta(s) + \int_0^t f(\varepsilon W(s)) \frac{W(s)}{s} ds,$$

which is the sum of a \mathcal{F}^W martingale, a time reversal of an \mathcal{H} martingale and a bounded variation term.

Proof. This is an immediate consequence of Theorem 4 and (2) since $W(t)/t$ is integrable on the interval $[0, T]$ and f is bounded. \square

3. PROOF OF THEOREM 4

Consider the time reversal \bar{W} of the Brownian Motion W . Note that the natural filtration of \bar{W} , $\mathcal{F}^{\bar{W}}$, is such that the increments of W in the interval $[T-t, T]$ are $\mathcal{F}_t^{\bar{W}}$ measurable. Further, it is easy to see that \bar{W} is a Brownian Motion adapted to its own filtration. Indeed, to see this it is enough to see that, for $0 \leq s < t \leq T$, the increment $\bar{W}(t) - \bar{W}(s) = W(T-t) - W(T-s)$ is a zero mean Gaussian random variable with variance $t-s$ independent of $\mathcal{F}_s^{\bar{W}}$.

Theorem [13, Theorem VI.3] applied to the Brownian Motion \bar{W} implies that the process

$$(6) \quad \begin{aligned} \beta(t) &= \bar{W}(t) - \int_0^t \frac{\bar{W}(T) - \bar{W}(s)}{T-s} ds \\ &= \bar{W}(t) + \int_0^t \frac{\hat{W}(s)}{T-s} ds \quad t \in [0, T], \end{aligned}$$

is a martingale with respect to \mathcal{G} . Moreover, from Hölder's inequality, it is easy to see that the integral in (6) is taken with respect to an integrable function and hence has finite variation. Then, it follows that $[\beta, \beta](t) = t$ and from Lévy's theorem β is a \mathcal{G} Brownian Motion. Hence, using (6) and Doob–Meyer decomposition theorem, we conclude that \bar{W} is a \mathcal{G} semimartingale with decomposition (3).

It remains to show that β is a \mathcal{H} Brownian Motion. Let us recall Jacod's Theorem [13, Theorem VI.2] which in this case implies that β is a Brownian Motion with respect to \mathcal{H} if $\beta(t)$ is independent of $W(T)$ for any $t \in [0, T]$. In order to show this independence, take $t \in [0, T]$ and note that $\beta(t)$ and $W(T)$ are jointly Gaussian centered with

$$\begin{aligned} \mathbb{E} W(T) \beta(t) &= \mathbb{E} ((W(T-t) - W(T)) W(T)) + \int_0^t \mathbb{E} \left(\frac{W(T-s)}{T-s} W(T) \right) ds \\ &= 0. \end{aligned}$$

Hence β is a Brownian Motion with respect to \mathcal{H} and \hat{W} is a semimartingale with respect to \mathcal{H} with decomposition (3).

Finally, since the increments of β are independent from $W(T)$, we can consider the SDE

$$(7) \quad d\hat{W}(t) = -\frac{\hat{W}(t)}{T-t} dt + d\beta(t), \quad \hat{W}(0) = W(T).$$

Let us show uniqueness for this equation. Suppose that $X(t)$ is another process that solves the same SDE with the same initial condition. Then, it is easy to see that $X(t) - \hat{W}(t) = C \cdot (T-t)$, for some random variable C and for any $t \in [0, T]$.

Hence, from $X(0) = \hat{W}(0)$, we conclude that $C = 0$ and the SDE (7) has a unique solution. Formula (4) follows by the variation of constants formula, and the theorem is proved.

Theorem 4 is the main element we need to propose our approximation scheme, which is the next step in the exposition.

4. SMALL NOISE ANALYSIS OF QUADRATIC COVARIATION.

In this section we study the quadratic covariation process $L_\varepsilon = [f(\varepsilon W), W]$. Recall the representation (5) given in Corollary 6. This will be one of the main ingredients in our proof.

Throughout this section, let $(n_\varepsilon)_{\varepsilon>0}$ be integers such that $n_\varepsilon \nearrow \infty$ as $\varepsilon \rightarrow 0$. Let us define $(\delta_\varepsilon)_{\varepsilon>0}$ by $\delta_\varepsilon = T/n_\varepsilon$, and observe that $\delta_\varepsilon \searrow 0$ as $\varepsilon \rightarrow 0$. The main result of this section is the following:

Theorem 7. *Let $q_\varepsilon = 2\sqrt{\delta_\varepsilon |\log \delta_\varepsilon|}$, and let $(\gamma_\varepsilon)_{\varepsilon>0}$ be such that $\gamma_\varepsilon \rightarrow 0$ and*

$$|\log \delta_\varepsilon| \frac{\text{osc}_f(\varepsilon q_\varepsilon)}{q_\varepsilon \gamma_\varepsilon} \rightarrow 0, \quad \varepsilon \rightarrow 0.$$

Then, there are positive constants K_1, K_2, K_3 and ε_0 such that

$$\mathbb{P} \left\{ \varepsilon^{-1} \sup_{t \leq T} |Q_\varepsilon(t)| > \gamma_\varepsilon \right\} \leq K_1 \gamma_\varepsilon^{-1} e^{-K_3 \gamma_\varepsilon^2 / \text{osc}_f(\varepsilon q_\varepsilon)^2} + K_2 \delta_\varepsilon, \quad \varepsilon \in (0, \varepsilon_0).$$

The idea of the proof is to start with the representation (1) and use Theorem 4 to prove an approximation to each integral by a sum of increments. The result will follow once we combine the approximating sum for each integral into one. We will devote the rest of this section to developing this idea.

4.1. Approximating Processes. Let P_ε be the partition of the interval $[0, T]$ given by points $0 = s_0 < \dots < s_{n_\varepsilon} = T$, where $s_i = i\delta_\varepsilon$, for $i = 0, \dots, n_\varepsilon$. Also, define the backward partition \hat{P}_ε to be the partition of $[0, T]$ given by points $0 = t_0 < \dots < t_{n_\varepsilon} = T$, where $t_i = T - s_{n_\varepsilon - i}$.

For an arbitrary process Y and times $s, t \in [0, T]$ let $\Delta_{t,s} Y = Y(t) - Y(s)$. Then, for $t \in [0, T]$ we introduce the following notation:

$$(8) \quad S_\varepsilon(t) = \int_0^t f(\varepsilon W(s)) dW(s),$$

$$(9) \quad \hat{S}_\varepsilon(t) = \int_{T-t}^T f(\varepsilon \hat{W}(s)) d\hat{W}(s),$$

$$(10) \quad J_\varepsilon(t) = \sum_{i=1}^{i(t)} f(\varepsilon W(s_{i-1})) \Delta_{s_i, s_{i-1}} W,$$

$$(11) \quad \hat{J}_\varepsilon(t) = \sum_{i=1}^{i(t)} f(\varepsilon W(s_i)) \Delta_{s_i, s_{i-1}} W,$$

where $i(t)$ is given by

$$i(t) = \min \{j \in [0, n_\varepsilon] \cap \mathbb{Z} : s_j \geq t\}.$$

The idea is to approximate each element S_ε and \hat{S}_ε with J_ε and \hat{J}_ε respectively, so we can approximate L_ε by $L_{\varepsilon, P_\varepsilon} = \hat{J}_\varepsilon - J_\varepsilon$. Note that since

$$f(\varepsilon W(s_i)) \Delta_{s_i, s_{i-1}} W = -f(\varepsilon \hat{W}(t_{n_\varepsilon - i})) \Delta_{t_{n_\varepsilon - i + 1}, t_{n_\varepsilon - i}} \hat{W},$$

after reordering the sum in (11), we can rewrite \hat{J}_ε as

$$(12) \quad \hat{J}_\varepsilon(t) = - \sum_{i=n_\varepsilon - i(t)}^{n_\varepsilon - 1} f(\varepsilon \hat{W}(t_i)) \Delta_{t_{i+1}, t_i} \hat{W},$$

which is an integral sum of the Itô integral \hat{S}_ε . We will use Theorem 4 to justify the application of martingale techniques to prove that J_ε approximates S_ε and that \hat{J}_ε approximates \hat{S}_ε .

Once we have an approximation of L_ε by $L_{\varepsilon, P_\varepsilon}$, we notice that

$$(13) \quad L_{\varepsilon, P_\varepsilon}(t) = \sum_{i=1}^{i(t)} \Delta_{s_i, s_{i-1}} (f(\varepsilon W)) \Delta_{s_i, s_{i-1}} W.$$

The differences in f in the above expression will be used to prove that $L_{\varepsilon, P_\varepsilon}(t)$ converges to 0 uniformly in probability and get the result.

We start with some preliminary results. The proofs will be postponed until Section 6 in order to keep the continuity of the paper. We state the next general lemma.

Lemma 8. *Let $(M_\varepsilon)_{\varepsilon > 0}$ be a family of martingales such that for every $\varepsilon > 0$, $M_\varepsilon(0) = 0$, the quadratic variation $\langle M_\varepsilon \rangle$ is absolutely continuous with respect to Lebesgue measure, and $\langle M_\varepsilon \rangle(T) \leq r_\varepsilon$. Then, for any $\delta > 0$,*

$$\mathbb{P} \left\{ \sup_{t \leq T} |M_\varepsilon(t)| > \delta \right\} < \sqrt{\frac{8r_\varepsilon}{\pi\delta^2}} e^{-\delta^2/(2r_\varepsilon)}.$$

We give a slight generalization of Levy's modulus of continuity lemma:

Lemma 9. *For a Brownian motion B , define the modulus of continuity with respect to partition P_ε by*

$$(14) \quad \delta_{B, \varepsilon} = \max_{i=1, \dots, n_\varepsilon} \sup_{s \in [s_{i-1}, s_i]} |\Delta_{s, s_{i-1}} B|.$$

Then, there is a constant $C > 0$ independent of $\varepsilon > 0$ such that for any $\delta > 0$

$$\mathbb{P} \{ \delta_{B, \varepsilon} > \delta \} \leq \frac{C}{\delta \sqrt{\delta_\varepsilon}} e^{-\delta^2/(2\delta_\varepsilon)}.$$

In particular, there is a $K_2 > 0$ such that

$$\mathbb{P} \{ \delta_{B, \varepsilon} > q_\varepsilon \} \leq K_2 \delta_\varepsilon, \quad \varepsilon > 0.$$

With these two results at hand we are ready to estimate $L_{\varepsilon, P_\varepsilon}$

Lemma 10. *There is a positive constant K such that for any $\delta > 0$ and $\varepsilon > 0$,*

$$\mathbb{P} \left\{ \sup_{t \in [0, T]} |L_{\varepsilon, P_\varepsilon}(t)| > \delta \right\} \leq \mathbb{P} \left\{ |\log \delta_\varepsilon| \text{osc}_f(\varepsilon q_\varepsilon) > \frac{q_\varepsilon \delta}{4T} \right\} + K_2 \delta_\varepsilon.$$

Of course, the probability in the r.h.s. is either 0 or 1, and the estimate is meaningful only if the inequality in the curly brackets is violated.

Proof. Let us start with the simple inequality

$$(15) \quad \sup_{t \in [0, T]} L_{\varepsilon, P_\varepsilon}(t) \leq \sum_{i=1}^{n_\varepsilon} |\Delta_{s_i, s_{i-1}} f(\varepsilon W)| |\Delta_{s_i, s_{i-1}} W|,$$

derived from (13). We estimate each term of the sum in the r.h.s. of (15).

From definition (14) it follows that

$$\max_{i=1, \dots, n_\varepsilon} |\Delta_{s_i, s_{i-1}} f(\varepsilon W)| \leq \text{osc}_f(\varepsilon \delta_{W, \varepsilon}).$$

Using this inequality and the definition of n_ε in (15), we see that

$$\begin{aligned} \sup_{t \in [0, T]} L_{\varepsilon, P_\varepsilon}(t) &\leq n_\varepsilon \text{osc}_f(\varepsilon \delta_{W, \varepsilon}) \delta_{W, \varepsilon} \\ &\leq T \delta_{W, \varepsilon} \text{osc}_f(\varepsilon \delta_{W, \varepsilon}) / \delta_\varepsilon. \end{aligned}$$

Hence for every $\delta > 0$ the inequalities

$$\begin{aligned} \mathbb{P} \left\{ \sup_{t \in [0, T]} |L_{\varepsilon, P_\varepsilon}| > \delta \right\} &\leq \mathbb{P} \{ \text{osc}_f(\varepsilon \delta_{W, \varepsilon}) \delta_{W, \varepsilon} > \delta_\varepsilon \delta / T, \delta_{W, \varepsilon} \leq q_\varepsilon \} + \mathbb{P} \{ \delta_{W, \varepsilon} > q_\varepsilon \} \\ (16) \quad &\leq \mathbb{P} \{ \text{osc}_f(\varepsilon q_\varepsilon) q_\varepsilon > \delta_\varepsilon \delta / T \} + \mathbb{P} \{ \delta_{W, \varepsilon} > q_\varepsilon \} \end{aligned}$$

hold. The second term in the r.h.s. of (16) can be bounded using Lemma 9, so we focus on the first term. For this notice that

$$\text{osc}_f(\varepsilon q_\varepsilon) q_\varepsilon / \delta_\varepsilon = 4 |\log \delta_\varepsilon| \text{osc}_f(\varepsilon q_\varepsilon) / q_\varepsilon,$$

which implies that

$$\mathbb{P} \{ \text{osc}_f(\varepsilon q_\varepsilon) q_\varepsilon > \delta_\varepsilon \delta / T \} \leq \mathbb{P} \{ |\log \delta_\varepsilon| \text{osc}_f(\varepsilon q_\varepsilon) > q_\varepsilon \delta / (4T) \}.$$

The result follows after combining this fact with (16) and Lemma 9. \square

4.2. Approximation of L_ε by $L_{\varepsilon, P_\varepsilon}$. We have shown that $L_{\varepsilon, P_\varepsilon}$ converges to 0. In order to prove the convergence of L_ε we need to prove that $L_{\varepsilon, P_\varepsilon}$ approximates L_ε .

In order to do so define

$$M_\varepsilon(t) := S_\varepsilon(t) - J_\varepsilon(t) + f(\varepsilon W(s_{i(t)-1})) \Delta_{s_{i(t)}, t} W,$$

and

$$\hat{M}_\varepsilon(t) := \hat{S}_\varepsilon(t) + \hat{J}_\varepsilon(t) + f(\varepsilon \hat{W}(t_{n_\varepsilon - i(t)})) \Delta_{T-t, n_\varepsilon - i(t)} \hat{W}.$$

Using (8), (10), and the definition of $i(t)$, we see that the process M_ε can be written as

$$M_\varepsilon(t) = \sum_{i=1}^{n_\varepsilon} \int_{s_{i-1} \wedge t}^{s_i \wedge t} \Delta_{s, s_{i-1}} f(\varepsilon W) dW(s).$$

Likewise, using (9), (11), (12) and the definition of the points t_i , we see that

$$\begin{aligned} \hat{M}_\varepsilon(t) &= \sum_{i=0}^{n_\varepsilon - 1} \int_{t_i \vee (T-t)}^{t_{i+1} \vee (T-t)} \Delta_{s, t_i} f(\varepsilon \hat{W}) d\hat{W}(s) \\ (17) \quad &= \sum_{i=0}^{n_\varepsilon - 1} \int_{t_i \vee (T-t)}^{t_{i+1} \vee (T-t)} \Delta_{s, t_i} f(\varepsilon \hat{W}) d\beta(s) - A_\varepsilon(t), \end{aligned}$$

where we defined

$$(18) \quad A_\varepsilon(t) = \sum_{i=0}^{n_\varepsilon-1} \int_{t_i \vee (T-t)}^{t_{i+1} \vee (T-t)} \Delta_{s, t_i} f(\varepsilon \hat{W}) \frac{\hat{W}(s)}{T-s} ds.$$

Notice that M_ε is a \mathcal{F}^W martingale and \hat{M}_ε is the time reversal of a \mathcal{F}^β semi-martingale. This is the main fact in the proof of the following Lemma:

Lemma 11. *There are positive constants K_1, K_2, K_3 and ε_0 such that for any $\delta > 0$,*

$$\mathbb{P} \left\{ \sup_{t \leq T} |\tilde{M}_\varepsilon(t)| > \delta \right\} \leq (K_1/\delta) e^{-K_3 \delta^2 / \text{osc}_f(\varepsilon q_\varepsilon)^2} + K_2 \delta_\varepsilon, \quad \varepsilon \in (0, \varepsilon_0).$$

Here \tilde{M}_ε can be either M_ε or \hat{M}_ε .

The following lemma will be used in the proof of Lemma 11. The proof is postponed until Section 6.

Lemma 12. *There are positive constants K_1, K_2, K_4 , and ε_0 such that for all $\delta > 0$,*

$$\mathbb{P} \left\{ \sup_{t \in (0, T)} |A_\varepsilon(t)| > \delta \right\} \leq (K_1/\delta) e^{-K_4 \delta^2 / \text{osc}_f(\varepsilon q_\varepsilon)^2} + K_2 \delta_\varepsilon, \quad \varepsilon \in (0, \varepsilon_0).$$

Proof of Lemma 11. Let us start with the proof for M_ε . As we said before, the process M_ε is a martingale with quadratic variation $\Gamma_\varepsilon = \langle M_\varepsilon \rangle$ given by

$$(19) \quad \Gamma_\varepsilon(t) = \sum_{i=1}^{n_\varepsilon} \int_{s_{i-1} \wedge t}^{s_i \wedge t} |\Delta_{s, s_{i-1}} f(\varepsilon W)|^2 ds.$$

In order to apply Lemma 8, we need to find a bound on the (random) function Γ_ε . In this case (14) implies that

$$\sup_{s \in [s_{i-1}, s_i]} |\Delta_{s, s_{i-1}} f(\varepsilon W)| \leq \text{osc}_f(\varepsilon \delta_{W, \varepsilon}),$$

for all $\varepsilon > 0$. Using this bound in (19) we see that

$$(20) \quad \Gamma_\varepsilon(T) \leq T \text{osc}_f(\varepsilon \delta_{W, \varepsilon})^2.$$

Lemma 8 implies that

$$(21) \quad \begin{aligned} \mathbb{P} \left\{ \sup_{t \leq T} |M_\varepsilon(t)| > \delta \right\} &\leq \mathbb{P} \left\{ \sup_{t \leq T} |M_\varepsilon(t)| > \delta, \Gamma_\varepsilon(T) \leq T \text{osc}_f(\varepsilon q_\varepsilon)^2 \right\} \\ &\quad + \mathbb{P} \left\{ \Gamma_\varepsilon(T) > T \text{osc}_f(\varepsilon q_\varepsilon)^2 \right\} \\ &\leq \sqrt{8T \frac{\text{osc}_f(\varepsilon q_\varepsilon)^2}{\pi \delta^2}} e^{-\delta^2 / (2T \text{osc}_f(\varepsilon q_\varepsilon)^2)} + \mathbb{P} \left\{ \Gamma_\varepsilon(T) > T \text{osc}_f(\varepsilon q_\varepsilon)^2 \right\}, \end{aligned}$$

for all $\varepsilon > 0$ small enough. It remains to estimate the second probability in (21). Using (20) it easily follows that for each $\varepsilon > 0$,

$$\begin{aligned} \mathbb{P} \left\{ \Gamma_\varepsilon(T) > T \text{osc}_f(\varepsilon q_\varepsilon)^2 \right\} &\leq \mathbb{P} \left\{ \text{osc}_f(\varepsilon \delta_{W, \varepsilon}) > \text{osc}_f(\varepsilon q_\varepsilon) \right\} \\ &\leq \mathbb{P} \left\{ \delta_{W, \varepsilon} > q_\varepsilon \right\}. \end{aligned}$$

Lemma 9 and (21) imply the desired estimate for M_ε .

To obtain the estimate on \hat{M}_ε , we notice that (17) and (18) imply

$$\begin{aligned}\hat{M}_\varepsilon(T-t) + A_\varepsilon(T-t) &= \sum_{i=0}^{n_\varepsilon-1} \int_{t_i \vee t}^{t_{i+1} \vee t} \Delta_{s,t_i} f(\varepsilon \hat{W}) d\beta(s) \\ &= \sum_{i=0}^{n_\varepsilon-1} \int_{t_i}^{t_{i+1}} \Delta_{s,t_i} f(\varepsilon \hat{W}) d\beta(s) - \sum_{i=0}^{n_\varepsilon-1} \int_{t_i \wedge t}^{t_{i+1} \wedge t} \Delta_{s,t_i} f(\varepsilon \hat{W}) d\beta(s).\end{aligned}$$

Then, it follows that

$$\begin{aligned}\sup_{t \leq T} |\hat{M}(t) - A_\varepsilon(t)| &= \sup_{t \leq T} |\hat{M}(T-t) - A_\varepsilon(T-t)| \\ &\leq \left| \sum_{i=0}^{n_\varepsilon-1} \int_{t_i}^{t_{i+1}} \Delta_{s,t_i} f(\varepsilon \hat{W}) d\beta(s) \right| + \sup_{t \leq T} \left| \sum_{i=0}^{n_\varepsilon-1} \int_{t_i \wedge t}^{t_{i+1} \wedge t} \Delta_{s,t_i} f(\varepsilon \hat{W}) d\beta(s) \right| \\ &\leq 2 \sup_{t \leq T} \left| \sum_{i=0}^{n_\varepsilon-1} \int_{t_i \wedge t}^{t_{i+1} \wedge t} \Delta_{s,t_i} f(\varepsilon \hat{W}) d\beta(s) \right|.\end{aligned}$$

Using this bound to proceed in the same way as we did for M_ε , we obtain that for any $\delta > 0$,

$$\mathbf{P} \left\{ \sup_{t \leq T} |\hat{M}(t) - A_\varepsilon(T)| > \delta \right\} \leq (K_1/\delta) e^{-K_2 \delta^2 / \text{osc}_f(\varepsilon q_\varepsilon)^2} + K_2 \delta_\varepsilon,$$

for all $\varepsilon > 0$ small enough. Since

$$\mathbf{P} \left\{ \sup_{t \leq T} |\hat{M}(t)| > \delta \right\} \leq \mathbf{P} \left\{ \sup_{t \leq T} |\hat{M}(t) - A_\varepsilon(t)| > \delta/2 \right\} + \mathbf{P} \left\{ \sup_{t \leq T} |A_\varepsilon(t)| > \delta/2 \right\},$$

the result follows from Lemma 12. \square

A consequence of Lemma 11 is the approximation of the quadratic covariation $L_\varepsilon = [f(\varepsilon W), W]$ by $L_{\varepsilon, P_\varepsilon}$, given in the following Lemma:

Lemma 13. *If $(\gamma_\varepsilon)_{\varepsilon>0}$ is such that $\gamma_\varepsilon \rightarrow 0$ and $\text{osc}_f(\varepsilon q_\varepsilon) q_\varepsilon \gamma_\varepsilon^{-1} \rightarrow 0$ as $\varepsilon \rightarrow 0$, then there are positive constants K_1, K_2, K_5 , and ε_0 such that*

$$\mathbf{P} \left\{ \sup_{t \leq T} |L_\varepsilon(t) - L_{\varepsilon, P_\varepsilon}(t)| > \gamma_\varepsilon \right\} \leq K_1 \gamma_\varepsilon^{-1} e^{-K_5 \gamma_\varepsilon^2 / \text{osc}_f(\varepsilon q_\varepsilon)^2} + K_2 \delta_\varepsilon, \quad \varepsilon \in (0, \varepsilon_0).$$

Proof. Let $(\gamma_\varepsilon)_{\varepsilon>0}$ be as in the statement of the Lemma. By the definition of M_ε and \hat{M}_ε , it follows that

$$(22) \quad |L_\varepsilon(t) - L_{\varepsilon, P_\varepsilon}(t)| \leq |M_\varepsilon| + |\hat{M}_\varepsilon| + |\Delta_{s_{i(t)}, s_{i(t)-1}} f(\varepsilon W) \Delta_{s_{i(t)}, s_{i(t)-1}} W|.$$

The result follows as a consequence of Lemmas 9 and 11. Indeed, since

$$|\Delta_{s_{i(t)}, s_{i(t)-1}} f(\varepsilon W) \Delta_{s_{i(t)}, s_{i(t)-1}} W| \leq \text{osc}_f(\varepsilon \delta_{W, \varepsilon}) \delta_{W, \varepsilon},$$

Lemma 9 implies

$$\begin{aligned}\mathbf{P} \{ |\Delta_{s_{i(t)}, s_{i(t)-1}} f(\varepsilon W) \Delta_{s_{i(t)}, s_{i(t)-1}} W| > \gamma_\varepsilon \} &\leq \mathbf{P} \{ \text{osc}_f(\varepsilon \delta_{W, \varepsilon}) \delta_{W, \varepsilon} > \gamma_\varepsilon, \delta_{W, \varepsilon} \leq q_\varepsilon \} \\ &\quad + \mathbf{P} \{ \delta_{W, \varepsilon} > q_\varepsilon \} \\ &\leq \mathbf{P} \{ \text{osc}_f(\varepsilon q_\varepsilon) q_\varepsilon > \gamma_\varepsilon \} + K_2 \delta_\varepsilon.\end{aligned}$$

Hence, there is a $\varepsilon_0 > 0$ such that

$$\mathbf{P} \{ |\Delta_{s_{i(t)}, s_{i(t)-1}} f(\varepsilon W) \Delta_{s_{i(t)}, s_{i(t)-1}} W| > \gamma_\varepsilon \} \leq K_2 \delta_\varepsilon, \quad \varepsilon \in (0, \varepsilon_0).$$

Using this bound and Lemma 11 in (22), we obtain

$$\mathbb{P} \left\{ \sup_{t \leq T} |L_\epsilon(t) - L_{\epsilon, P_\epsilon}(t)| > \gamma_\epsilon \right\} \leq K_1 \gamma_\epsilon^{-1} e^{-K_5 \gamma_\epsilon^2 / \text{osc}_f(\epsilon q_\epsilon)^2} + K_2 \delta_\epsilon, \quad \epsilon \in (0, \epsilon_0).$$

The proof is finished. \square

5. PROOF OF THEOREMS 1, 2 AND 7

Proof of Theorem 7. The result is a consequence of Lemmas 10 and 13. Indeed, if $(\gamma_\epsilon)_{\epsilon > 0}$ is as in the statement of the Theorem, it is immediate to see that

$$(23) \quad \mathbb{P} \left\{ \epsilon^{-1} \sup_{t \leq T} |Q_\epsilon(t)| > \gamma_\epsilon \right\} \leq \mathbb{P} \left\{ \sup_{t \leq T} |L_\epsilon(t) - L_{\epsilon, P_\epsilon}(t)| > \gamma_\epsilon / 2 \right\} \\ + \mathbb{P} \left\{ \sup_{t \in [0, T]} |L_{\epsilon, P_\epsilon}(t)| > \gamma_\epsilon / 2 \right\}.$$

The result will follow by applying Lemmas 10 and 13 to the two terms in r.h.s. of (23).

First, note that

$$\eta_\epsilon = |\log \delta_\epsilon| \frac{\text{osc}_f(\epsilon q_\epsilon)}{q_\epsilon \gamma_\epsilon} \rightarrow 0, \quad \epsilon \rightarrow 0,$$

implies that $\text{osc}_f(\epsilon q_\epsilon) q_\epsilon \gamma_\epsilon^{-1} = 4 \eta_\epsilon \delta_\epsilon \rightarrow 0$, as $\epsilon \rightarrow 0$. Hence, from Lemma 13 we get that for some positive constants K'_1, K_2, K'_5 and ϵ'_0

$$(24) \quad \mathbb{P} \left\{ \sup_{t \leq T} |L_\epsilon(t) - L_{\epsilon, P_\epsilon}(t)| > \gamma_\epsilon / 2 \right\} \leq K'_1 \gamma_\epsilon^{-1} e^{-K'_5 \gamma_\epsilon^2 / \text{osc}_f(\epsilon q_\epsilon)^2} + K_2 \delta_\epsilon,$$

for all $\epsilon \in (0, \epsilon_0)$. Likewise, since $\eta_\epsilon \rightarrow 0$ as $\epsilon \rightarrow 0$, Lemma 10 implies that for some positive constants ϵ_1 and K ,

$$(25) \quad \mathbb{P} \left\{ \sup_{t \in [0, T]} |L_{\epsilon, P_\epsilon}(t)| > \gamma_\epsilon / 2 \right\} \leq K_2 \delta_\epsilon, \quad \epsilon \in (0, \epsilon_1).$$

The result follows by using (24) and (25) in (23). \square

Proof of Theorem 1. The proof is a consequence of Theorem 7. Indeed, let us find a family $(\delta_\epsilon)_{\epsilon > 0}$ such that $\delta_\epsilon \rightarrow 0$ and

$$\lim_{\epsilon \rightarrow 0} |\log \delta_\epsilon| \frac{\text{osc}_f(\epsilon q_\epsilon)}{q_\epsilon} = 0.$$

Let $A(\delta_\epsilon, \epsilon) = |\log \delta_\epsilon| \text{osc}_f(\epsilon q_\epsilon) q_\epsilon^{-1}$. A straightforward calculation gives

$$A(\delta_\epsilon, \epsilon) \leq C_f \epsilon^\alpha \delta_\epsilon^{(\alpha-1)/2} |\log \delta_\epsilon|^{(\alpha+1)/2}.$$

Let $\mu \in (\gamma, \alpha)$ and take $\delta_\epsilon = \epsilon^{2(\alpha-\mu)/(1-\alpha)}$. Then, $A(\epsilon^{2(\alpha-\mu)/(1-\alpha)}, \epsilon) \leq \hat{A}(\epsilon)$, where $\hat{A}(\epsilon)$ is given by

$$\hat{A}(\epsilon) = C_{\alpha, f} \epsilon^\mu |\log \epsilon|^{(\alpha+1)/2},$$

for some constant $C_{\alpha,f} > 0$ independent of $\varepsilon > 0$. So, we can use this δ_ε in Theorem 7 to get that

$$(26) \quad \mathbb{P} \left\{ \varepsilon^{-1} \sup_{t \leq T} |Q_\varepsilon(t)| > \delta \right\} \leq K_1 \delta^{-1} \exp \left\{ -C_0 \frac{(\delta \varepsilon^{-\alpha(1-\mu)/(1-\alpha)})^2}{|\log \varepsilon|^\alpha} \right\} \\ + K_2 \varepsilon^{2(\alpha-\mu)/(1-\alpha)},$$

for all $\varepsilon > 0$ small enough and constants $K_1, K_2, C_0 > 0$ independent of $\varepsilon > 0$ and $\delta > 0$.

Theorem 7 actually implies that inequality (26) remains true as long as $\hat{A}(\varepsilon)/\delta \rightarrow 0$, as $\varepsilon \rightarrow 0$. So, since $\gamma \in (0, \mu)$, we can substitute $\varepsilon^\gamma \delta$ for δ in (26) to get that

$$(27) \quad \mathbb{P} \left\{ \varepsilon^{-(1+\gamma)} \sup_{t \leq T} |Q_\varepsilon(t)| > \delta \right\} \leq K_1 \delta^{-1} \varepsilon^{-\gamma} \exp \left\{ -C_0 \frac{\delta^2 \varepsilon^{2(-(\alpha-\gamma)+\alpha(\mu-\gamma))/(1-\alpha)}}{|\log \varepsilon|^\alpha} \right\} \\ + K_2 \varepsilon^{2(\alpha-\mu)/(1-\alpha)}.$$

Since $\alpha \in (0, 1)$ and $\mu < \alpha$, we have

$$\alpha(\mu - \gamma) < \alpha(\alpha - \gamma) < \alpha - \gamma.$$

Using this fact in (27) we get that

$$\mathbb{P} \left\{ \varepsilon^{-(1+\gamma)} \sup_{t \leq T} |Q_\varepsilon(t)| > \delta \right\} \leq K_3 \varepsilon^{2(\alpha-\mu)/(1-\alpha)},$$

for some $K_3 > 0$, any $\delta > 0$, and all $\varepsilon > 0$ small enough. The result is proved. \square

Proof of Theorem 2. The proof follows the same steps as the proof of Theorem 1. The first step is to follow Theorem 7 by finding a family $(\delta_\varepsilon)_{\varepsilon>0}$ such that $\delta_\varepsilon \rightarrow 0$ and

$$\lim_{\varepsilon \rightarrow 0} |\log \delta_\varepsilon| \varepsilon = 0.$$

Given $\gamma \in (0, 1)$, we propose $\delta_\varepsilon = e^{-\varepsilon^{-(1-\mu)}}$, for $\mu \in (\gamma, 1)$. In this case, $|\log \delta_\varepsilon| \varepsilon = \varepsilon^\mu$, so Theorem 7 implies that

$$\mathbb{P} \left\{ \varepsilon^{-1} \sup_{t \leq T} |Q_\varepsilon(t)| > \delta \right\} \leq K_1 \delta^{-1} \exp \left\{ -K_4 \delta^2 \varepsilon^{-(1+\mu)} e^{\varepsilon^{-(1-\mu)}} \right\} \\ + K_2 e^{-\varepsilon^{-(1-\mu)}}.$$

As in the proof of Theorem 1, we can substitute $\delta \varepsilon^\gamma$ instead of δ in the last inequality. We can finish the proof by extracting the leading term in the resulting estimate. \square

6. ADDITIONAL PROOFS

Proof of Lemma 8. For each $\epsilon > 0$, we use the representation of martingales as time changed Brownian Motion [9, Theorem 3.4.2] to see that $M_\epsilon = B(\langle M_\epsilon \rangle)$ in distribution in the space of continuous functions, for some Brownian Motion B (see [9, Theorem 3.4.2]). Therefore,

$$\mathbb{P} \left\{ \sup_{t \leq T} |M_\epsilon(t)| > \delta \right\} \leq \mathbb{P} \left\{ \sup_{t \leq r_\epsilon} |B(t)| > \delta \right\}.$$

Now the symmetry of B , reflection principle [9, Section 2.6], and Brownian scaling (self-similarity) imply that

$$\begin{aligned} \mathbb{P} \left\{ \sup_{t \leq r_\epsilon} |B(t)| > \delta \right\} &= \mathbb{P} \left\{ \sup_{t \leq r_\epsilon} \max\{B(t), -B(t)\} > \delta \right\} \\ &\leq 2\mathbb{P} \left\{ \sup_{t \leq r_\epsilon} B(t) > \delta \right\} \\ &\leq 4\mathbb{P} \{B(r_\epsilon) > \delta\} \\ &= 4\mathbb{P} \{\sqrt{r_\epsilon}B(1) > \delta\}. \end{aligned}$$

The result follows by a standard Gaussian Tail estimate. \square

Proof of Lemma 9. Fix $\delta > 0$ and note that

$$(28) \quad \mathbb{P} \{ \delta_{B,\epsilon} > \delta \} \leq \sum_{i=1}^{n_\epsilon} \mathbb{P} \left\{ \sup_{s \in (s_{i-1}, s_i)} |\Delta_{s, s_{i-1}} B| > \delta \right\}.$$

We bound each of the probabilities in this sum. Since the process $\Delta_{s, s_{i-1}} B$ is equal in distribution, on the space of continuous functions, to a Brownian Motion itself up to a time shift, we can use reflection principle [9, Theorem 2.9.25] and standard Gaussian bounds to get

$$\begin{aligned} \mathbb{P} \left\{ \sup_{s \in (s_{i-1}, s_i)} |\Delta_{s, s_{i-1}} B| > \delta \right\} &\leq 4\mathbb{P} \{B(\delta_\epsilon) > \delta\} \\ &\leq \delta^{-1} \sqrt{\frac{8\delta_\epsilon}{\pi}} e^{-\delta^2/2\delta_\epsilon}. \end{aligned}$$

Substituting this expression in (28) and using the fact that $n_\epsilon \leq 2T/\delta_\epsilon$, we see that there is a constant $C > 0$ independent of $\epsilon > 0$ such that for any $\delta > 0$

$$\mathbb{P} \{ \delta_{B,\epsilon} > \delta \} \leq \frac{C}{\delta \sqrt{\delta_\epsilon}} e^{-\delta^2/(2\delta_\epsilon)}$$

as expected.

To prove the second part, use $\delta = q_\epsilon = 2\sqrt{-\delta_\epsilon \log \delta_\epsilon}$ in the last expression to get that

$$\begin{aligned} \mathbb{P} \{ \delta_{B,\epsilon} > q_\epsilon \} &\leq \frac{C}{2\delta_\epsilon \sqrt{-\log \delta_\epsilon}} e^{2 \log \delta_\epsilon} \\ &= \frac{C\delta_\epsilon}{2\sqrt{-\log \delta_\epsilon}} \\ &\leq K_2 \delta_\epsilon, \end{aligned}$$

for some constant $K_2 > 0$. Hence the result follows. \square

Proof of Lemma 12. We start with a basic inequality

$$\begin{aligned} \sup_{t \in [0, T]} |A_\varepsilon(t)| &\leq \sum_{i=0}^{n_\varepsilon-1} \int_{s_i}^{s_{i+1}} |\Delta_{s, s_i} f(\varepsilon \hat{W})| \frac{|\hat{W}(s)|}{T-s} ds \\ &\leq 2\sqrt{T} \text{osc}_f(\varepsilon \delta_{W, \varepsilon}) \sup_{s \leq T} \frac{|\hat{W}(s)|}{\sqrt{T-s}} \\ &\leq 2\sqrt{T} \text{osc}_f(\varepsilon \delta_{W, \varepsilon}) \sup_{s \leq T} \frac{|W(s)|}{\sqrt{s}}. \end{aligned}$$

It implies that

$$\begin{aligned} \mathbf{P} \left\{ \sup_{t \in [0, T]} |A_\varepsilon(t)| > \delta \right\} &\leq \mathbf{P} \left\{ \sup_{s \leq T} \frac{|W(s)|}{\sqrt{s}} > \frac{\delta}{2 \text{osc}_f(\varepsilon q_\varepsilon) \sqrt{T}} \right\} + \mathbf{P} \{ \delta_{W, \varepsilon} > q_\varepsilon \} \\ &\leq \mathbf{P} \left\{ \sup_{s \leq T} \frac{|W(s)|}{\sqrt{s}} > \frac{\delta}{2 \text{osc}_f(\varepsilon q_\varepsilon) \sqrt{T}} \right\} + K_1 \delta_\varepsilon, \end{aligned}$$

for some constant $K_1 > 0$ independent of $\delta > 0$ and $\varepsilon > 0$. To finish the proof, we need to study the tail probability of the random variable $A = \sup_{s \leq T} |W(s)|/\sqrt{s}$.

In order to study the tail decay of the random variable A , note that, due to the symmetry of Brownian Motion,

$$\mathbf{P} \{ A > \delta \} \leq 2\mathbf{P} \left\{ \sup_{t \leq T} \frac{W(t)}{\sqrt{t}} > \delta \right\}.$$

So it is sufficient to focus on the tail probabilities of the random variable $N = \sup_{t \leq T} (W(t)/\sqrt{t})$, which is the supremum of a Gaussian process.

Equip the interval $[0, T]$ with the metric ρ given by

$$\begin{aligned} \rho(s, t)^2 &= \mathbf{E} \left(\frac{W(s)}{\sqrt{s}} - \frac{W(t)}{\sqrt{t}} \right)^2 \\ &= 2 \left(1 - \sqrt{\frac{s \wedge t}{s \vee t}} \right), \quad s, t \in [0, T]. \end{aligned}$$

We denote by $B_\theta(t) \subset [0, T]$ the ρ -ball of radius $\theta > 0$ centered at $t \in [0, T]$. Let H_θ be the minimum number of balls of radius θ needed in order to cover $[0, T]$. According to [10][Section 14, Theorem 1], if

$$(29) \quad \int_0^{\sigma/2} \sqrt{|\log H_\theta|} d\theta < \infty,$$

with $\sigma = \sup_{t \in [0, T]} \text{var}(W(t)/\sqrt{t}) = 1$, then $\mathbf{E}N < \infty$. Then, it is standard to see [10][Corollary 2, Section 14] that there is a $\zeta_0 > \mathbf{E}N$, such that for any $\zeta > \zeta_0$

$$(30) \quad \mathbf{P} \{ |N - \mathbf{E}N| > \zeta \} \leq C e^{-\zeta^2/2} / \zeta,$$

for some universal constant $C > 0$.

In our situation, if the integral in (29) is finite, this will be enough to finish the proof. Indeed, assuming (30), there is an $\varepsilon_0 > 0$ such that

$$\begin{aligned} \mathbb{P} \left\{ \sup_{s \leq T} \frac{|W(s)|}{\sqrt{s}} > \frac{\delta}{2 \text{osc}_f(\varepsilon q_\varepsilon) \sqrt{T}} \right\} &\leq 2 \mathbb{P} \left\{ N > \frac{\delta}{2 \text{osc}_f(\varepsilon q_\varepsilon) \sqrt{T}} \right\} \\ &\leq 2 \mathbb{P} \left\{ N - \mathbb{E}N > \frac{\delta}{2 \text{osc}_f(\varepsilon q_\varepsilon) \sqrt{T}} - \mathbb{E}N \right\} \\ &\leq C_1 \left(\frac{\delta}{2 \sqrt{T} \text{osc}_f(\varepsilon q_\varepsilon)} - \mathbb{E}N \right)^{-1} \exp \left\{ -C_2 \left(\frac{\delta}{2 \sqrt{T} \text{osc}_f(\varepsilon q_\varepsilon)} - \mathbb{E}N \right)^2 \right\} \\ &\leq (C_1/\delta) \exp \left\{ -C_3 \frac{\delta^2}{\text{osc}_f(\varepsilon q_\varepsilon)^2} \right\}, \end{aligned}$$

for some constants $C_1, C_2, C_3 > 0$ independent of ε and δ , and all $\varepsilon \in (0, \varepsilon_0)$. Hence we just need to show that the integral (29) is finite.

We are going to give an estimate of H_θ , $\theta \in (0, 1/2)$. Suppose $0 \leq s < t \leq T$, then $s \in B_\theta(t)$ if and only if

$$\sqrt{s} \geq \sqrt{t}(1 - \theta^2/2).$$

Therefore, if s and t belong to the same ball of radius $\theta \in (0, 1/2)$, then

$$|t - s| \leq T\theta^2.$$

Hence, $H_\theta \leq 2/\theta^2$, and $\sqrt{|\log H_\theta|}$ is integrable on the interval $[0, 1/2]$, which implies our claim. \square

REFERENCES

- [1] Yuri Bakhtin. Noisy heteroclinic networks. *Probability Theory and Related Fields*, 150:1–42, 2011. 10.1007/s00440-010-0264-0.
- [2] Blandine Bérard Bergery and Pierre Vallois. Approximation via regularization of the local time of semimartingales and Brownian motion. *Stochastic Process. Appl.*, 118(11):2058–2070, 2008.
- [3] N. Bouleau. Formules de changement de variables. *Ann. Inst. H. Poincaré Probab. Statist.*, 20(2):133–145, 1984.
- [4] Nicolas Bouleau and Marc Yor. Sur la variation quadratique des temps locaux de certaines semimartingales. *C. R. Acad. Sci. Paris Sér. I Math.*, 292(9):491–494, 1981.
- [5] Nathalie Eisenbaum. Integration with respect to local time. *Potential Anal.*, 13(4):303–328, 2000.
- [6] Nathalie Eisenbaum. On Itô’s formula of Föllmer and Protter. In *Séminaire de Probabilités, XXXV*, volume 1755 of *Lecture Notes in Math.*, pages 390–395. Springer, Berlin, 2001.
- [7] Nathalie Eisenbaum. Local time-space calculus for reversible semimartingales. In *Séminaire de Probabilités XL*, volume 1899 of *Lecture Notes in Math.*, pages 137–146. Springer, Berlin, 2007.
- [8] Hans Föllmer, Philip Protter, and Albert N. Shiryaev. Quadratic covariation and an extension of Itô’s formula. *Bernoulli*, 1(1-2):149–169, 1995.
- [9] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1988.
- [10] M. A. Lifshits. *Gaussian random functions*, volume 322 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 1995.
- [11] Sergio Angel Almada Monter and Yuri Bakhtin. Normal forms approach to diffusion near hyperbolic equilibria. *Nonlinearity*, 24(6):1883, 2011.
- [12] Lawrence Perko. *Differential equations and dynamical systems*, volume 7 of *Texts in Applied Mathematics*. Springer-Verlag, New York, third edition, 2001.

- [13] Philip E. Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, second edition, 2004. Stochastic Modelling and Applied Probability.
- [14] F. Russo and P. Vallois. Itô formula for C^1 -functions of semimartingales. *Probab. Theory Related Fields*, 104(1):27–41, 1996.
- [15] Francesco Russo and Pierre Vallois. The generalized covariation process and Itô formula. *Stochastic Process. Appl.*, 59(1):81–104, 1995.
- [16] Francesco Russo and Pierre Vallois. Elements of stochastic calculus via regularization. In *Séminaire de Probabilités XL*, volume 1899 of *Lecture Notes in Math.*, pages 147–185. Springer, Berlin, 2007.